Reliability of two common classifications for Legg-Calvé-Perthes disease if using MRI: Pediatric orthopedists vs. radiologists

Serkan Erkus1, Onder Kalenderer2, Ali Turgut3, Dilek Oncel4
1Specialist in Orthopaedics and Traumatology, Izmir, Türkiye
2Izmir Tınaztepe University, Faculty of Health Sciences, Department of Occupational Therapy, Izmir, Türkiye
3Tepecik Training and Research Hospital, Department of Orthopaedics and Traumatology, Izmir, Türkiye
4Tepecik Training and Research Hospital, Department of Radiology, Izmir, Türkiye

Received 08 September 2022; Accepted 30 September 2022

Abstract
In patients with Legg-Calvé-Perthes (LCP), the Catterall and Herring classifications based on standard hip radiographs are frequently used. In this study, the compatibility of classification systems in hip radiographs and MR images of LCP patients was examined among specialist working in the departments of Orthopedics and Radiology, and the reliability was investigated. In this study, which included 37 Perthes patients, two separate presentations were prepared in which the radiological imaging of the patients was determined by drawing lots. Six experienced specialists were selected as observers, three of whom were radiologist and three of pediatric orthopedists. Each presentations were sent to the observers one month apart. They were asked to categorize visual data. Data were evaluated with percent agreement (PA) and Gwet’s AC1 method. In the Catterall classification, agreement among pediatric orthopedists decreased in MRI examination, while an increase was observed among radiologists. In the Herring classification, while pediatric orthopedists remained at a similar level, radiologists’ inter-observer reliability increased. When the intra-observer reliability was examined, it was observed that the radiologists in the Catterall classification and the orthopedists in the Herring classification were more stable in the evaluations made with MRI. The usability of both classification systems, which provide prognostic prediction, in MRI examinations is a crucial unknown. It is fruitful that both radiologists and pediatric orthopedists work together in novel classification systems to be created using MRI.

Keywords: Hip radiograph, MRI, legg-calvé-perthes disease, reliability

Introduction
Accurate interpretation of the femoracetabular relationship is momentous in the treatment and follow-up of Legg-Calvé-Perthes (LCP) disease, an osteochondrosis affecting the proximal epiphysis of the femur [1]. Moreover, the concept of containment of the femoral head within the acetabulum guides the option of suitable treatment [1-3]. A preliminary assessment of this relationship can be made with radiography. This prediction can be further explored with MRI [4,5]. Thus, the coverage of the femoral head by the acetabulum can be demonstrated in more details.

There are two classification systems that are associated with radiographic findings and the severity of the disease and are frequently used in the fragmentation stage: Catterall [6] and Herring [1] classification systems. The common feature of these two classifications is the use of radiological findings reflecting the degree of femoral proximal involvement. Although MRI is an expensive method and cannot be used frequently due to the need for sedation in pediatric patients to ensure immobility, it is an advantageous imaging method that can show hip joint compatibility in all aspects. In the literature, it is possible to discover studies reporting findings that can predict prognosis using MRI [4,7-10]. However, there is still no defined and validated classification method for MRI. The ease of use, inter-observer reproducibility, and prognostic predictive value of the classification system are directly related to the validity of the system [11]. The aim of this study is to investigate the usability of Catterall and Herring classification systems in MRI and to examine the intra- and inter-observer reliability by healthcare professionals from different...
fields of expertise.

Materials and Methods

Radiographic imaging of 30 patients with LCPD were included in this cross-sectional descriptive study. No patient identification data was included in the visual data. The study was exempt from the requirement of informed consent because of the retrospective nature of the study and the anonymous data. This study was carried out following the Declaration of Helsinki's principles. This cross-sectional study was conducted following the approval of the local ethics committee (Decision no: 2019/6, Date 16.03.2019).

Visual data included radiographs and magnetic resolution imaging (MRI) of both hips. The MRI production date was similar or close to the radiographs'. Hip radiographs were obtained with neutral A-P and frog-leg position. For MRI, a 1.5 Tesla scanner (Avanto, Siemens AG Medical Solutions, Erlangen, Germany) was used. The following sequences were obtained from PACS: T1-weighted coronal spin echo (TR/TE 500/20-ms, FOV: 19-cm, Slice thickness: 4-mm, Intersection gap: 0.4-mm) and T2-weighted coronal gradient echo (TR/TE 600/27-ms, FOV: 21-cm, Slice thickness: 4-mm, Intersection gap: 0.4-mm) images were obtained (FOV: field of volume, TR/TE: repetition time/echo time).

Firstly, two different interpretations were created. At the beginning of each, theoretical and visual information about the two classification systems were included. The order of data was randomly changed for second. The first included X-ray, and the second included MRI. Three orthopedists with at least 10 years of experience in pediatric orthopedics and 3 radiologists with at least 10 years of experience in pediatric musculoskeletal imaging were included as observers. Two interpretations were sent to all observers to be evaluated one month apart. Washout period was accepted as a month. The observers were asked to classify the visual data of each radiographs or MRI by the Catterall and Herring classification.

Statistics were analyzed using the R software package v.3.4.2 (The R Foundation, Vienna, Austria). Percentage agreement (PA) and inter-observer agreement and intra-observer reproducibility were calculated using Gwet’s AC1 [12]. The interpretation of the data was performed according to the Landis and Koch kappa benchmarks (<0: poor agreement; 0-0.20: slight agreement; 0.21-0.40: fair agreement; 0.41-0.60: moderate agreement; 0.61-0.80: substantial agreement; and 0.81-1.00: almost perfect agreement) [13].

Results

All of the 30 patients with LCPD included in the study were in the fragmentation stage. Reliability for first reading was found to be fair for both of Catterall and Herring classification systems among pediatric orthopedists. In the evaluation made with MRI (second reading), these values were again found to be fair. Among radiologists, it was seen that the values of substantial for Catterall classification and moderate for Herring classification, which were determined in the first evaluation, resulted as substantial and almost perfect in the second evaluation, respectively.

When the intra-observer reliability values were examined, it was determined that the most significant change for pediatric orthopedists was in Catterall classification (PA: 41.2%) in separate evaluations made with different imaging methods. For radiologists, a similar change was observed in Catterall and Herring classification (PA: 46.1%, 48.0%). Detailed results of the study are shown in Tables 1 and 2.

Table 1. Inter-observer reliability

<table>
<thead>
<tr>
<th>Classification</th>
<th>Pediatric Orthopaedists</th>
<th>Radiologists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PA (%)</td>
<td>Gwet’s AC1</td>
</tr>
<tr>
<td>Catterall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set - 1</td>
<td>50.0</td>
<td>0.38</td>
</tr>
<tr>
<td>Set - 2</td>
<td>42.2</td>
<td>0.29</td>
</tr>
<tr>
<td>Herring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set - 1</td>
<td>48.0</td>
<td>0.36</td>
</tr>
<tr>
<td>Set - 2</td>
<td>48.0</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Table 2. Intra-observer reliability

<table>
<thead>
<tr>
<th>Classification</th>
<th>Catterall</th>
<th>Herring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PA (%)</td>
<td>Gwet’s AC1</td>
</tr>
<tr>
<td>Pediatric Orthopaedists – 1</td>
<td>35.3</td>
<td>0.15</td>
</tr>
<tr>
<td>Pediatric Orthopaedists – 2</td>
<td>47.0</td>
<td>0.30</td>
</tr>
<tr>
<td>Pediatric Orthopaedists – 3</td>
<td>41.2</td>
<td>0.22</td>
</tr>
<tr>
<td>Overall</td>
<td>41.2</td>
<td>0.22</td>
</tr>
<tr>
<td>Radiologists – 1</td>
<td>47.0</td>
<td>0.30</td>
</tr>
<tr>
<td>Radiologists – 2</td>
<td>50.0</td>
<td>0.34</td>
</tr>
<tr>
<td>Radiologists – 3</td>
<td>41.2</td>
<td>0.22</td>
</tr>
<tr>
<td>Overall</td>
<td>46.1</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Discussion

The most important implication of this study was that there were intra-observer and inter-observer reliability differences in groups of Orthopedists and also Radiologists in the radiological classifications used in Perthes disease. When MR imaging was accepted as a parameter in the classification system, it was found to increase the reliabilities. We believe that in the future, radiologists should be included in the study groups in classifications based on MR.

Commonly used classification systems for the early and fragmentation stages of Perthes disease are the Catterall and Herring classification. Catterall classification is made by considering the degree of femoral head involvement [6]. The lateral pillar–Herring classification is associated with the height of the lateral pillar of the capital femoral epiphysis [1] and is also the most predictive at the fragmentation stage [14]. Both classification systems are applied with radiography of pelvis. These classifications technically contain some advantages and disadvantages. Variable intra- and interobserver agreement rates have been reported in both the Catterall and Herring classifications...
in the early and middle stages of the disease [14-22]. Classification systems have a key role in constituting a universal language among all user physicians that indicates the severity and character of the disease. Its reproducibility, ease of use, high predictiveness in terms of prognosis affect the importance and prevalence of using the classification system. In order to use systems for the purpose and to create a universal language, reliability studies are observed objectively are needed. Therefore, in the literature review, there are many intra- and inter-observer compatibility studies for these two classification systems. MRI evaluations, which have entered daily practice, have begun to take their place in decision-making in the diagnosis and treatment of Perthes disease. To the best of our knowledge, studies on the validity of these classification methods in MRI are scarce. There are studies on findings that may have other prognostic importance by examining MRI findings [7-10,23]. In addition to illuminating this deficiency, this study has crucial implications in terms of reflecting the views of radiologists other than orthopedists.

When the reproducibility of the Catterall classification evaluations are examined in the literature, it is seen that generally low levels of interobserver reliability are detected [14,17,24]. However, it is possible to come across studies showing that the increase in professional experience in the field of pediatric orthopedics increases these rates [18]. On the contrary, in a reliability study conducted by fragmentation phase patients, it was reported by Kalenderer et al. that agreement was at fair levels and experience was not an advantage in compliance [19]. In this study, while inter-observer reliability was fair among pediatric orthopedists, it was observed to be substantial among radiologists. In the MRI-containing readings (Set-2), the agreement levels did not change in both groups, but the value decreased among orthopedists (from 0.38 to 0.29), while a slight increase (from 0.68 to 0.71) was observed among the radiologists. Moreover, there was 50% agreement at first reading among pediatric orthopedists versus 73.5% for radiologists. However, these rates changed to 42.2% and 76.5%, respectively, at the second reading. While staging of disease with MRI among pediatric orthopedists likely varied, an increase in mildly for agreement was observed among radiologists. The Catterall classification is a system related to the affected surface area of the femoral head. In Turkey, where the study was conducted, reporting may have resulted in accurate measurements. Thus, they could have correctly evaluated the patients in the Staging system.

In this study, T1-weighted coronal spin echo and T2-weighted coronal gradient echo sequences were used in the MRI method. However, nowadays, studies on MRI examinations obtained in different sequences with or without contrast are at the forefront with their advantages [23]. Due to the retrospective nature of the study, MRI scans could not be altered. This is the most critical limitation of this study. In addition, no causality was examined to the observers as to why their opinions changed, and the detected changes were evaluated subjectively.

Conclusion

In LCP disease, the involvement of the femoral head can be distinguished in the early period by using MRI. Some prognostic factors have been identified in MRI studies. A novel classification method can be created with studies that combine this advanced knowledge with radiological examination findings that have proven prognostic importance. Orthopedists can predict disease prognosis and eventual changes, while radiologists have an advantage in interpreting technical examinations such as MRI. For this reason, it will be important to establish classification systems with high validity and reproducibility in cooperation with orthopedists and radiologists.

Conflict of interests

The authors declare that there is no conflict of interest in the study.

Financial Disclosure

The authors declare that they have received no financial support for the study.

Ethical approval

This study was carried out following the Declaration of Helsinki's principles. This cross-sectional study was conducted following the approval of the local ethics committee (Decision no: 2019/6, Date 16.03.2019).

References

4. Mak MS, Teh J. Magnetic resonance imaging of the hip: anatomy and


